SUPPLEMENTAL INFORMATION FOR

Small Molecules Enhance CRISPR Genome Editing in Pluripotent Stem Cells

Chen Yu, Yanxia Liu, Tianhua Ma, Kai Liu, Shaohua Xu, Yu Zhang, Honglei Liu, Marie La Russa, Min Xie, Sheng Ding, Lei S. Qi

Supplemental Experimental Procedures

Construction of sgRNA vectors and DNA templates

To clone sgRNA mCherry vectors, the optimized sgRNA expression vector (pSLQ1651, Addgene catalog no. 51024) was linearized via double digestion with BstXI and XhoI, and gel purified. New sgRNA sequences were PCR amplified from pSLQ1651 using different forward primers (see below) and a common reverse primer (sgRNA.R), digested with BstXI and XhoI, gel purified, and ligated to the linearized pSLQ1651 vector.

sgNanog.F: GGAGA ACCAC CTTGT TGGCG TAAGT CTCAT ATTTC ACCGT TTAAG AGCTA TGCTG GAAAC AGCA

sgSOD1.F: GTATC CCTTG GAGAA CCACC TTGTT GGTCG CCCTT CAGCA CGCAC AGTTT AAGAG CTATG CTGGA AACAG CA

sgRNA.R: CTAGT ACTCG AGAAA AAAAG CACCG ACTCG GTGCC AC

To clone a single Cas9-sgRNA expressing vector, the pX330 (Addgene catalog no. 42230) expression vector expressing Cas9 and sgRNA was linearized with BbsI digestion, and gel purified. A pair of oligos for each targeting site were phosphorylated, annealed, and ligated to the linearized pX330.

sgGFP-1.F: CACCG CATCA CCTTC ACCCT CTCCA

sgGFP-1.R: AAACT GGAGA GGGTG AAGGT GATGC

sgGFP-2.F: CACCG CGTGC TGAAG TCAAG TTTGA

sgGFP-2.R: AAACT CAAAC TTGAC TTCAG CACGC

sgGFP-3.F: CACCGTCGACAGGTAATGGTTGTC

sgGFP-3.R: AAACG ACAAC CATTA CCTGT CGAC

sgACTA2.F: CACCG CGGTG GACAA TGGAA GGCC

sgACTA2.R: AAACG GCCTT CCATT GTCCA CCGC

The p2A-NLS-sfGFP template of *Nanog* was assembled from four DNA fragments, a 5' homology arm, a p2A-NLS_{X2}-sfGFP cassette, a 3' homology arm, and a modified pUC19 backbone vector, using Gibson Assembly Master Mix (New England Biolabs). Both 5' and 3' homology arms were PCR amplified from the genomic DNA extracted from mouse ES cells. The sequences of p2A and two copies of NLS were added to the upstream of sfGFP coding sequence by PCR amplification. The backbone vector was linearized by digestion with PmeI and ZraI. All DNA fragments were gel purified before the Gibson assembly reaction.

Cell culture, electroporation, and flow cytometry analysis

The E14 mouse ES cells were maintained in N2B27 medium (50% Neurobasal, 50% Dulbecco modified Eagle medium/Ham's nutrient mixture F12, 0.5% NEAA, 0.5% Sodium Pyruvate, 0.5% GlutaMax, 0.5% N2, 1% B27, 0.1mM β-mercaptoethanol and 0.05 g/L bovine albumin fraction V; all from Invitrogen) supplemented with LIF and 2i in gelatin-coated plates.

For electroporation, 3×10^6 cells were electroporated using the Nucleofector Kit for Mouse Embryonic Stem Cells (Amaxa) with program A-030. For insertion experiments, 2.5 μg pX330 (Cas9), 2.5 μg sgNanog and 15 μg template (Nanog-p2A-NLS-sfGFP) were used. For sfGFP deletion experiments, 20 μg pX330 containing desired sgRNA was used. All plasmids were maxiprepped using the Endofree Maxiprep Kit (Qiagen). Cells post electroporation were counted with trypan blue, seeded to Matrigel-coated plates in LIF-containing ESGRO-2i medium (Millipore), and cultured for 3 days. At day 3, cells were analyzed using the BD FACSCalibur platform.

Human ES cell-derived neural stem cells were cultured in N2B27 medium supplemented with 3 μM of CHIR99021 and 1 μM of A-83-01. Human fibroblasts (CRL-2097) and HeLa cells

were cultured in Dulbecco modified Eagle medium supplemented with 10% FBS (Gibco). K562 cells were cultured in RPMI medium supplemented with 10% FBS. HUVECs were culture using Endothelial Cell Growth Media Kit (Lonza). For DNA insertion at the *ACTA2* locus, 1×10^7 cells were electroporated with 5 μg pX330-sgACTA2 and 15 μg template using the Neon Transfection System (Life Technologies). The programs used were: 1,300 V, 10 ms, and 3 pulses for human ES cell-derived neural stem cells; 1,500 V, 30 ms, and 1 pulse for fibroblasts; 1,005 V, 35 ms, and 2 pulses for HeLa; 1,450 V, 10 ms, and 3 pulses for K562; and 1,350 V, 30ms, and 1 pulse for HUVEC. At day 3, fluorescence was analyzed using the BD FACSCalibur platform.

SOD1 SNP editing in human iPS cells

The human induced pluripotent stem (iPS) cells (hiPSC-O#1, Zhu et al., 2010) were cultured in mTeSR1 (STEMCELL Technologies) in Geltrex coated 6-well plates. Three hours prior electroporation, cells were moved to fresh mTeSR1 medium supplemented with 1 μM ROCK inhibitor (thiazovivin). Established method was used for the delivery of the Cas9 vector, sgSOD1 mCherry vector and the 200-nt ssODN template (sequence: 5'-GTGCT GGTTT GCGTC GTAGT CTCCT GCAGC GTCTG GGGTT TCCGT TGCAG TCCTC GGAAC CAGGA CCTCG GCGTG GCCTA GCGAG TTATG GCGAC GAAGG TCGTG TGCGT GCTGA AGGGC GACGG CCCAG TGCAG GGCAT CATCA ATTTC GAGCA GAAGG CAAGG GCTGG GACGG AGGCT TGTTT GCGAG GCCGC TCCCA-3') (Ding et al., 2013a; Ding et al., 2013b). Briefly, 1×10^7 cells were electroporated with a mixture of 15 μg Cas9 vector, 15 μg sgSOD1 mCherry vector with or without (no template control) 30 μg ssODN template using the BioRad Gene Pulser. Cells were then recovered in mTeSR1 medium supplemented with 1 μM ROCK inhibitor with or without L755507 for 48 hours after electroporation. The mCherry positive cells were collected by Fluorescence Activated Cell Sorting (FACS) into 6-well plates and culture for 5 days before genome DNA preparation using PureLink Genomic DNA Mini Kit (Life

Technologies). Genomic DNA was PCR amplified with Herculase II Fusion DNA polymerase (Agilent) using two primers flanking the homology arms (forward primer sequence: AAAGT GCCAC CTGAC AGGTC TGGCC TATAA AGTAG TCGCG; reverse primer sequence: AGCTG GAGAC CGTTT GACCC GCTCC TAGCA AAGGT). PCR products were purified using NucleoSpin Gel and PCR Cleanup Kit (Macherey-Nagel). The two primers contained extra 15-bp regions that allowed efficient subcloning onto a modified pUC19 vector using the In-Fusion HD Cloning Plus kit (Clontech). The cloning products were transformed into DH5α E. coli competent cells and grew on LB agar plates with Carbenicillin (Sigma). After overnight culture, we randomly picked 96, 288, and 192 colonies for no template, DMSO and L755507 samples, respectively. All E. coli colonies were minipreped and sequencing verified to detect the mutation sequences (QuintaraBio). The A4V allele mutant frequency is calculated as (# of A4V transformants)/(total # of bacterial transformants). The indel allele frequency is calculated as (# of indel transformants)/(total # of bacterial transformants). The allele that contained both A4V mutation and another indel was simply counted as an indel allele.

Sequencing of long template insertion at Nanog and ACTA2

For long template insertion at *Nanog* or *ACTA2* loci, genomic DNA from 1×10^6 cells were isolated and purified with PureLink Genomic DNA Mini Kit (Life Technologies). For sequencing, genomic DNA was PCR amplified with Herculase II Fusion DNA polymerase (Agilent) with a pair of primers outside homology arms. PCR products were purified and subcloned to a backbone vector (pUC19) using In-Fusion cloning for sequencing. The following PCR primers were used:

Nanog.F: AAAGT GCCAC CTGAC ATTCT TCTAC CAGTC CCAAA CAAAA GCTCTC

Nanog.R: AGCTG GAGAC CGTTT AGCAA ATGTC AATCC CAAAG TTGGG AG

ACTA2.F: AAAGT GCCAC CTGAC CTGGT TAGCC AGTTT TCAC TGTTC TCTGT

ACTA2.R: AGCTG GAGAC CGTTT GCATT TTGGA AAGTC AAGAG GAGAG AATTGC

For p2A-NLS_{x2}-sfGFP insertion, a primer (sequence: GCATG ACTTT TTCAA GAGTG CCA) that bound within sfGFP was used to confirm correct insertion.

Deep sequencing of Nanog-sfGFP knockout

For deep sequencing, the Nanog-sfGFP locus was PCR amplified and purified. Adapters and barcodes were added to amplicon by PCR. The DNA fragments were sequenced on a MiSeq (illumina) with MiSeq Reagent Kit v3 (150 cycles) following the manufacturer's instructions.

Nanog-sfGFP.F: ACACG TTCAG AGTTC TACAG TCCGA CGATC GACGG GACCT ACAAG ACGCG

Nanog-sfGFP.R: ACACG TTCAG AGTTC TACAG TCCGA CGATC GACGG GACCT ACAAG ACGCG

5' adapter primer: AATGA TACGG CGACC ACCGA GATCT ACACG TTCAG AGTTC TACAG TCCGA

3' barcode primers: CAAGC AGAAG ACGGC ATACG AGATA AACAG TGTGA CTGGA GTTCC TTGGC ACCCG AGAAT TCCA; CAAGC AGAAG ACGGC ATACG AGATA AACCC CGTGA CTGGA GTTCC TTGGC ACCCG AGAAT TCCA; CAAGC AGAAG ACGGC ATACG AGATA AACGG CGTGA CTGGA GTTCC TTGGC ACCCG AGAAT TCCA

Small molecule compound library and screening

Sigma LOPAC library (1280 compounds), Tocriscreen library (1120 compounds), and part of Spectrum Collection library (1760 compounds) were screened. For screening, 50 nL of each compound was added into each well of Matrigel-coated 384-well plates containing 20 μL ESGRO-2i medium. After electroporation, 2,000 cells in 70 uL ESGRO-2i medium were seeded to the 384-well plates. After 3 days culture, cells were fixed, stained with DAPI, and imaged using IN Cell analyzer (GE). The numbers of DAPI-positive nuclei and DAPI/GFP double-

positive nuclei were counted by IN cell analyzer. The ratio of double-positive nuclei and DAPI-positive nuclei was calculated and plotted from high to low as shown in Figure 1D. Extreme outliers were individually examined and excluded if the results were due to severe cell death.

Generation of a clonal mouse ES cell line carrying monoallelic sfGFP insertion at the Nanog locus

The E14 mouse ES cells electroporated with a template plasmid (p2A-NLS_{x2}-sfGFP) were cultured for 3 days and dissociated into single cells with Accutase (Life Technologies). Single GFP-positive cells were sorted and seeded to each wells of a Matrigel-coated 96-well plate with the FACS Aria II (BD). 7 days after sorting, clonal GFP-positive colonies were expanded in 6-well plates. A rabbit polyclonal antibody (abcam) was used for immunofluorescence staining of Nanog.

Toxicity assay

Cells were treated with small molecules during the first 24 hours after electroporation. Cell number was counted at day 3 post electroporation. Cell viability was measured by the MTS assay (Promega) following manufacturer's instructions.

SUPPLEMENTAL FIGURE LEGENDS

Figure S1, related to Figure 1. (A) A fluorescence histogram of mouse ES cells transfected with Cas9, sgNanog, and/or a control template containing p2A-sfGFP without the homology arms (HAs). (B) A scheme of the high-throughput chemical screening platform. (C) Characterization of GFP insertion efficiency at the *Nanog* locus in mouse ES cells with different treatment windows of four small molecules. (D) Cell number was counted at day 3 after post electroporation. Cells were treated with small molecules at the first 24 hours. (E) Cell viability was measured by the MTS assay (Promega). Absorbance at 490 nm was normalized to E14 cells. In Figures S1C-S1E, error bars represent the standard deviation of three biological replicates.

Figure S2, related to Figure 2. (A) A scheme of generating a clonal mouse ES cell line carrying a monoallelic sfGFP insertion at the *Nanog* locus. Two sets of primer binding sites are shown by arrows. One primer set (#1) binds to the sequences outside of the homology arms, and the other primer set (#2) contains a forward primer binding to the sfGFP sequence and a reverse primer binding outside of the 3' homology arm. (B) A gel picture showing validation of single allele tagging using two sets of primers. (C) Immunofluorescence of Oct4 and Sox2 of E14 cells treated with small molecules after 10 passages. Cells were treated with small molecules for the first 24 hours after splitting. (D) Flow cytometry analysis of Nanog of E14 cells treated with small molecules. (E) Microscopic images of Nanog-sfGFP ES cells electroporated with different sgRNAs. (F) Microscopic images of Nanog-sfGFP mouse ES cells electroporated with sgGFP-1 in the presence of DMSO, L755507 (5 μM), or AZT (1 μM). (G) Microscopic images of Nanog-sfGFP mouse ES cells treated with AZT for 10 passages. Cells were treated with small molecules during the first 24 hours after each splitting. For S2C, S2E-S2G, scale bars, 50 μm.

SUPPLEMENTAL TABLE

Table S1, related to Figure 2. Deep sequencing analysis of sfGFP targeting by sgGFP-2.

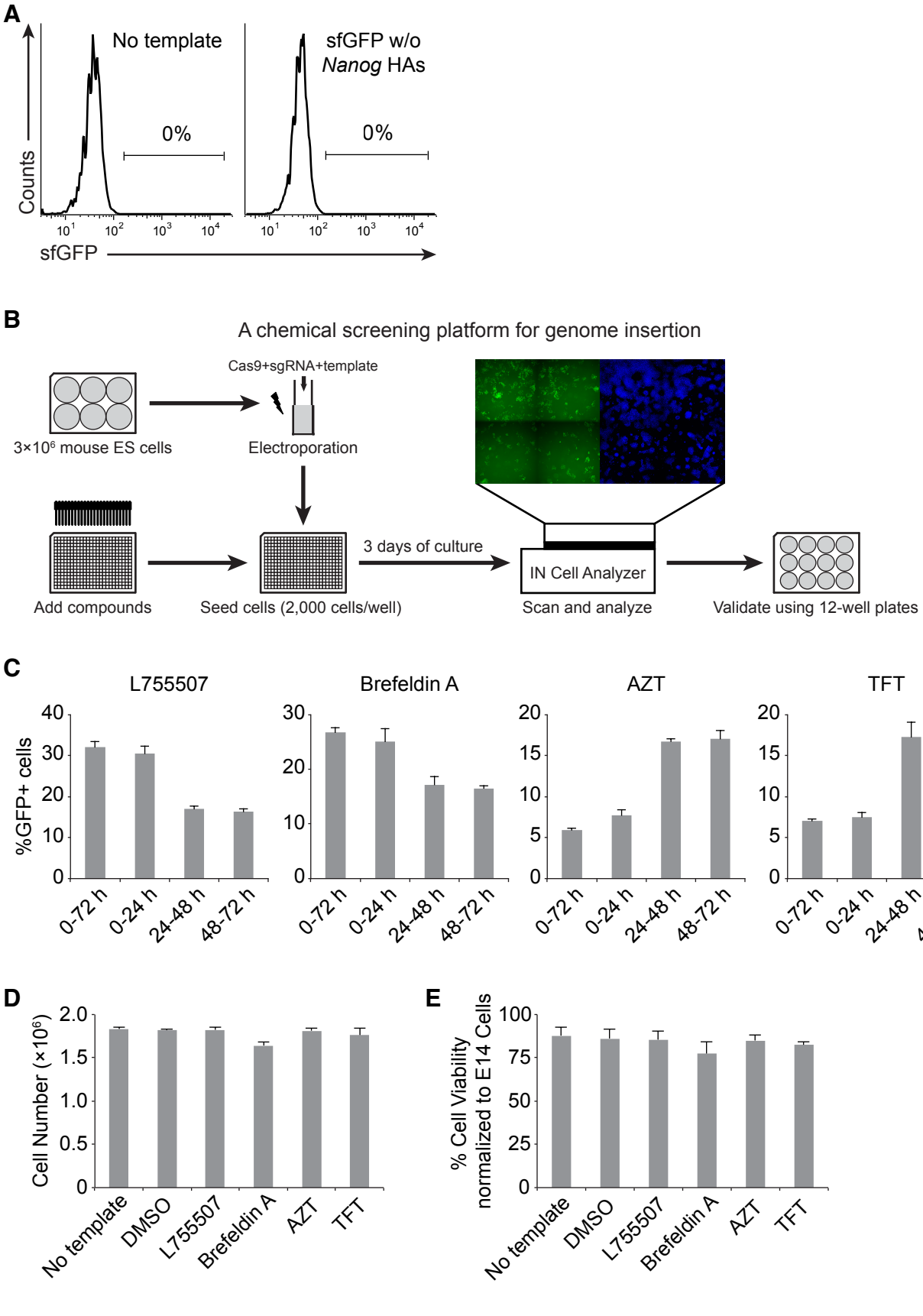
	DMSO	L755507	AZT
Indel mutation/ total reads	24,998/49,211	21,313/57,202	49,067/83,075
Indel %	50.80%	47.74%	59.06%

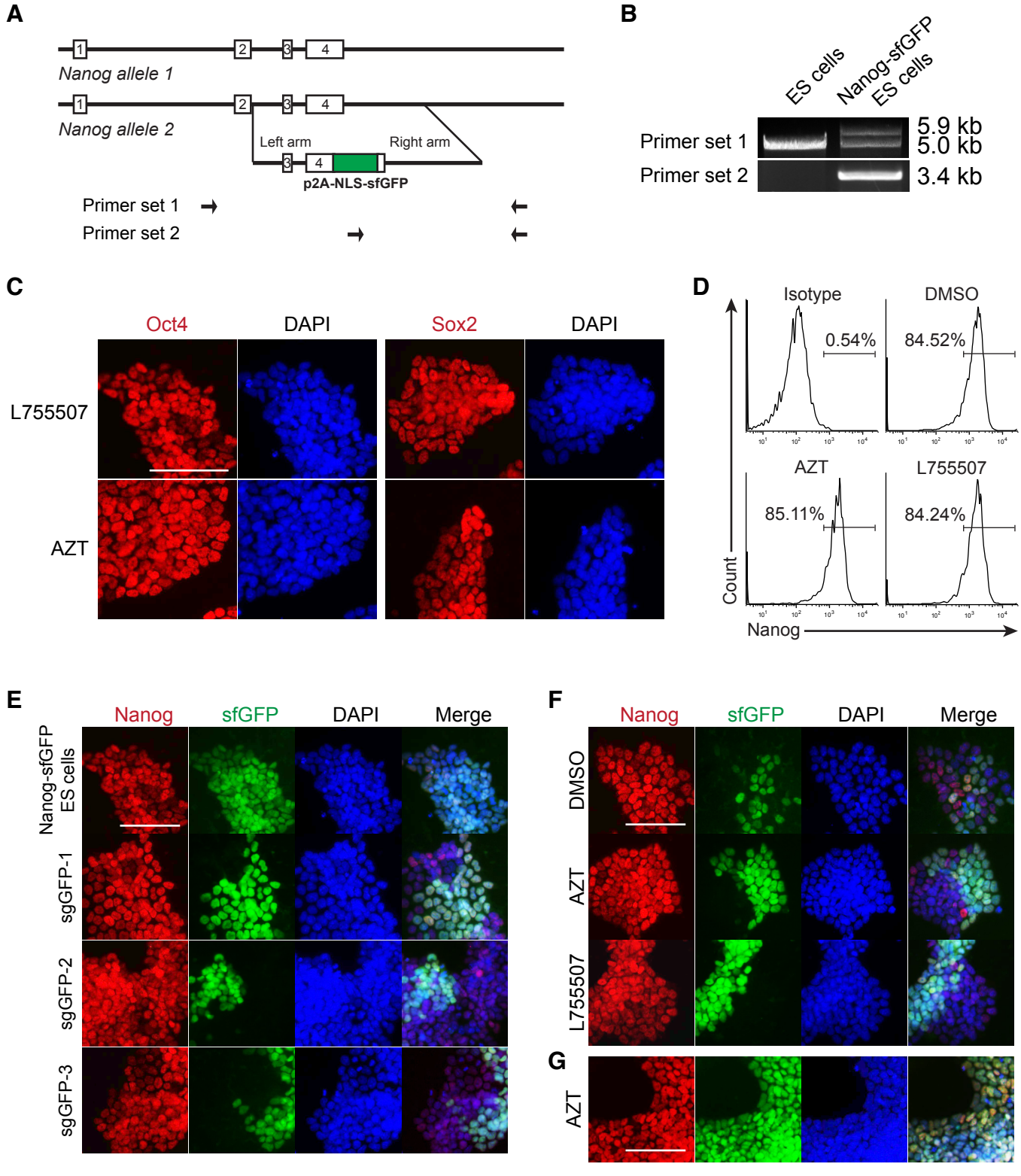
SUPPLEMENTAL REFERENCES

Ding, Q., Lee, Y.K., Schaefer, E.A., Peters, D.T., Veres, A., Kim, K., Kuperwasser, N., Motola, D.L., Meissner, T.B., Hendriks, W.T., et al. (2013a). A TALEN genome-editing system for generating human stem cell-based disease models. Cell stem cell 12, 238-251.

Ding, Q., Regan, S.N., Xia, Y., Oostrom, L.A., Cowan, C.A., and Musunuru, K. (2013b). Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. Cell stem cell 12, 393-394.

Zhu, S., Li, W., Zhou, H., Wei, W., Ambasudhan, R., Lin, T., Kim, J., Zhang, K., and Ding, S. (2010). Reprogramming of human primary somatic cells by OCT4 and chemical compounds. Cell stem cell 7, 651-655.


A
No template
sfGFP w/o
Nanog HAs
0%
0%
Counts
sfGFP
B
A chemical screening platform for genome insertion
Cas9+sgRNA+template
3×10⁶ mouse ES cells
Electroporation
3 days of culture
IN Cell Analyzer
Add compounds
Seed cells (2,000 cells/well)
Scan and analyze
Validate using 12-well plates
C
L755507
Brefeldin A
AZT
TFT
%GFP+ cells
0-72 h
0-24 h
24-48 h
48-72 h
D
Cell Number (×10⁶)
No template
DMSO
L755507
Brefeldin A
AZT
TFT
E
% Cell Viability
normalized to E14 Cells
No template
DMSO
L755507
Brefeldin A
AZT
TFT


A
Nanog allele 1
Nanog allele 2
Left arm
Right arm
p2A-NLS-sfGFP
Primer set 1
Primer set 2
B
ES cells
Nanog-sfGFP ES cells
Primer set 1
Primer set 2
5.9 kb
5.0 kb
3.4 kb
C
Oct4
DAPI
Sox2
DAPI
L755507
AZT
D
Isotype
DMSO
0.54%
84.52%
AZT
L755507
85.11%
84.24%
Count
Nanog
E
Nanog
sfGFP
DAPI
Merge
Nanog-sfGFP ES cells
sgGFP-1
sgGFP-2
sgGFP-3
F
Nanog
sfGFP
DAPI
Merge
DMSO
AZT
L755507
G
AZT